
**Nucleotide**

[PubMed](#)
[Nucleotide](#)
[Protein](#)
[Genome](#)
[Structure](#)
[PMC](#)
[Taxonomy](#)
[GAMMA](#)
[Books](#)

[My NCBI](#)
[\[Sign In\]](#)
[\[Register\]](#)

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full) Show 5 Send to

Range: from 1828653 to 1830242

Show whole sequence

☐ Reverse complemented strand For☒ 1: NC\_006461. Reports *Thermus thermophilus*...[gi:55979969]

Links

Comment Features Sequence

**LOCUS** NC\_006461 1590 bp DNA linear BCT 03-DEC-2005  
**DEFINITION** *Thermus thermophilus* HB8, complete genome.  
**ACCESSION** NC\_006461 REGION: 1828653..1830242  
**VERSION** NC\_006461.1 GI:55979969  
**PROJECT** GenomeProject:13202  
**KEYWORDS** .  
**SOURCE** *Thermus thermophilus* HB8  
**ORGANISM** *Thermus thermophilus* HB8  
 Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
 Thermus.  
**REFERENCE** 1  
**AUTHORS** Masui, R., Kurokawa, K., Nakagawa, N., Tokunaga, F., Koyama, Y.,  
 Shibata, T., Oshima, T., Yokoyama, S., Yasunaga, T. and Kuramitsu, S.  
**TITLE** Complete genome sequence of *Thermus thermophilus* HB8  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 1590)  
**CONSTRM** NCBI Genome Project  
**TITLE** Direct Submission  
**JOURNAL** Submitted (20-NOV-2004) National Center for Biotechnology  
 Information, NIH, Bethesda, MD 20894, USA  
**REFERENCE** 3 (bases 1 to 1590)  
**AUTHORS** Masui, R., Kurokawa, K., Nakagawa, N., Tokunaga, F., Koyama, Y.,  
 Shibata, T., Oshima, T., Yokoyama, S., Yasunaga, T. and Kuramitsu, S.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (11-NOV-2004) Graduate School of Information Science,  
 Nara Institute of Science and Technology, 8916-5, Takayamacho,  
 Ikoma, Nara 630-0192, Japan  
**COMMENT** PROVISIONAL REFSEQ: This record has not yet been subject to final  
 NCBI review. The reference sequence was derived from AP008226.  
 COMPLETENESS: full length.  
**FEATURES** Location/Qualifiers  
 source 1..1590  
 /organism="Thermus thermophilus HB8"  
 /mol\_type="genomic DNA"  
 /strain="HB8"  
 /db\_xref="taxon:300852"  
 gene 1..1590  
 /locus\_tag="TTHA1952"  
 /db\_xref="GeneID:3166520"  
 CDS 1..1590  
 /locus\_tag="TTHA1952"  
 /inference="non-experimental evidence, no additional  
 details recorded"  
 /note="similar to GB:AAS61930.1 percent identity 99 in 529  
 aa"  
 /codon\_start=1  
 /transl\_table=1

```

/product="DNA polymerase III holoenzyme tau subunit"
/protein_id="YE_145218.1"
/db_xref="GI:55981921"
/db_xref="GeneID:3168520"
/translation="MSALYRRFRPI.TFQEVVQGHVKEFLLEATREGSLAQAYLFSGCP
RGVGTGTTTARLLAMAVGCGQGEDFFCGVOPHCQAVCGGAHFDVVEIDARSNNSVEDVRE
LREKRIHLAPLSAPKRVFILDFAHMLSKAFNALLKTLEZPPHVLVFFATTEPERMFP
TILSRTOHFFFRKLTETEEIAFLKRLRIEAVGSPAEERERALLMARLADGALDAESLLE
RFLLESGPLTRNEVERALGLPFRALAEIRASLARGKTAEALGLARLHYGEGYAPRSE
VSGLLEVFEGLYRAFGLASTFLPAPPQALIAANTALDEAMERLARKSDALSLEVAL
EAGRALAAEALPQPTGAPSPFVCPKPFSPAPPEPRPFEPADLERWRFALEALRPTL
RAFVREARFVREGQLCLAFPEDKAFHYRKASEQKARILLPLAQAHTGVSEVVLVLEG
KSLSPRFPAPPPPEAPAFPGPEEEVEAEAEAEAEAEALRRVRLGGRLVNVRRP
RTRAEPRPEEPLSQDEIGGTGI"

```

## OPIGIN

```

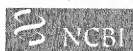
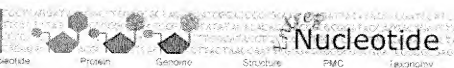
1  gtagagggccc  tctaccggcg  cttccggccc  ctcacattcc  aggaggtggt  ggggcaggag
61  cagctgaagg  agccctctcc  caaggccatc  cgggaggagg  ggctcgccca  ggctctacctc
121  ttctccggcg  ccaggggcgt  gggcaagacc  accacggcga  ggctctctgc  catggcggtg
181  gggcgccagg  ggaagagacc  cccttgccgg  gctgcccacc  actgccaggc  ggtgcagagg
241  cggcgccacc  cggacgtggt  ggagattgac  gcgcggagca  accactccgt  ggaggacgtg
301  cgggagctga  ggaagaggat  ccacctcgcc  cccctctctg  cccccaggaa  ggtctctatc
361  ctggacgagg  cccactgtgt  ctccaaaagc  gccctcaacg  cccctctcaa  gacctctggg
421  gagcccccg  cccacgtctc  ctttgtcttc  gccacccacc  agcccgagag  gatgcccccc
481  accatctctc  ccgcacccca  gcacttcgcc  ttccgcccgc  tcacggagga  ggagatcgcc
541  tttaagctcc  ggcgcacctc  ggaggccgtg  gggcgggagg  cggaggagga  ggccctctctc
601  ctctctggcc  cccctggcga  cggggccctt  agggacggcg  aaagccctct  ggagcctctc
661  ctctctctgg  aaggccccc  cccccggagg  gaggtgagac  gcgcctctag  cctccccccc
721  agggaggccc  tggccgagat  cgcgcctctc  ctcgcgagg  ggaacacggc  ggagggccgtg
781  ggctctggcc  ggccgctcta  cggggaagg  taagcccgca  ggaacctggt  ctccggccctt
841  ttggagggtg  tccgggaagg  cctctacgcc  gccctgggcc  tcgcgggaac  cccctctccc
901  gcccccggcc  aggcctctat  cgcgcgcatg  accgccttgg  acgagggcat  ggagcgcctc
961  gcccgccgct  ccgacgcctt  aagcctggag  gtggccctcc  tggaggcggg  aaggggccctg
1021  gccgcggagg  cctcgcccca  gccacggg  gctccctccc  caggagtggt  ccccaagccg
1081  gaaagccccc  cggccccgga  acccccagg  cccgaggagg  cgcgcgacct  cggggagcgg
1141  tggcgggcct  tctctgaggg  cctcaggccc  accctacggg  cctctgtcgg  gaggcccccg
1201  ccggaggctc  gggaaaggcca  gctctgcctc  gctttccccc  aggaacaggc  cttccactac
1261  cgaaggccct  cggaaacgaa  ggcgaggtcc  ctccccctgg  cccagggcca  tttcggggtg
1321  gaggagggtg  tctctgctct  ggaggagaaa  aaaaaaagcc  tgagcccaag  gcccccgctc
1381  gccccacctc  ctgaagcgcc  cgcacccccg  ggccctcccg  aaggaggagt  agaggcggag
1441  gaaagcggg  aggaagcccc  ggaggaggcc  ttgaggcggg  tggctccgct  cctggggggg
1501  cgggtgctct  ggggtcggg  gccaggacc  cgggagggcg  cggaggagga  accctgagc
1561  caagacgaga  taggggggtac  tggatatata

```

//


Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Aug 15 2006 13:27:38






**Nucleotide**

[My NCBI](#) [\[Sign In\]](#) [\[Register\]](#)

[Nucleotide](#) [Protein](#) [Genome](#) [Structure](#) [PMC](#) [Technology](#) [OMIM](#) [Books](#)

Search **Nucleotide**  for

Limits  Preview/Index History Clipboard Details

Display **GenBank(Full)**  Show **5**  Send to 

Range: from **1502604** to **1504193**  ☐ Reverse complemented strand [Fe](#)

☒ **I: AE017221. Reports** *Thermus thermophilus*...[gi:46197919]

[Links](#)

**Comment** **Features** **Sequence**

**LOCUS** AE017221 1590 bp DNA linear BCT 13-DEC-2005  
**DEFINITION** *Thermus thermophilus* HB27, complete genome.  
**ACCESSION** AE017221 REGION: 1502604..1504193  
**VERSION** AE017221.1 GI:46197919  
**KEYWORDS** .  
**SOURCE** *Thermus thermophilus* HB27  
**ORGANISM** *Thermus thermophilus* HB27  
 Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae; Thermus.  
**REFERENCE** 1 (bases 1 to 1590)  
**AUTHORS** Henne, A., Brueggemann, H., Raasch, C., Wiezer, A., Hartsch, T., Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R., Jacobi, C., Starkuviene, V., Schlenczek, S., Dencker, S., Huber, R., Klenk, H.-P., Overbeek, R., Kramer, W., Merkl, R., Gottschalk, G. and Fritz, H.-J.  
**TITLE** The genome sequence of the extreme thermophile *Thermus thermophilus*  
**JOURNAL** Nat. Biotechnol. 22 (5), 547-553 (2004)  
**PUBMED** 15064768  
**REFERENCE** 2 (bases 1 to 1590)  
**AUTHORS** Henne, A., Brueggemann, H., Raasch, C., Wiezer, A., Hartsch, T., Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R., Jacobi, C., Starkuviene, V., Schlenczek, S., Dencker, S., Huber, R., Klenk, H.-P., Overbeek, R., Kramer, W., Merkl, R., Gottschalk, G. and Fritz, H.-J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (03-FEB-2004) Goettingen Genomics Laboratory, Institute for Microbiology and Genetics, Grisebachstr. 8, Goettingen D-37077, Germany  
**COMMENT** On or before Dec 13, 2005 this sequence version replaced gi:46197930, gi:46196221, gi:46196498, gi:46196775, gi:46197055, gi:46197378, gi:46197659.  
**FEATURES**  
**source** Location/Qualifiers  
 1..1590  
 /organism="Thermus thermophilus HB27"  
 /mol\_type="genomic DNA"  
 /strain="HB27"  
 /db\_xref="taxon:262124"  
**gene** 1..1590  
 /locus\_tag="TT\_C1588"  
 /old\_locus\_tag="TTC1588"  
**CDS** 1..1590  
 /locus\_tag="TT\_C1588"  
 /old\_locus\_tag="TTC1588"  
 /EC\_number="2.7.7.7"  
 /codon\_start=1  
 /transl\_table=11  
 /product="DNA polymerase III holoenzyme tau subunit"

```

/protein_id="AA581930.1"
/db_xref="GI:46197817"
/translation="MSALYRFRPLTFQEVVGQEHVKEPLLKAIKREGPLAQAYLFSGP
RGVSKTTTARLLAMAVGCGQGEDPPCGVCPHCQAVQGAHPDVVETDAASNNVVEQVRE
LRERIHLAPLSAPRQVFIIDENHMLSKSAFNALLKILKEPPPHVLFVFATTEPERMPP
TILSKTQHPRFRRLTEFEIAFKLRRILKAVGREAESEALLLARIADGALRDAESILE
RFLLEGFPTPKVEVERALGLPPREALAEIASLARGKTASALGLARRLYGFGYAPRSL
VSGLLEVFREGLYAATGLAGTFLPAPPQALIAAMTALDEAMERLARSSDALSLLEVAL
ERGRAIAAEALPQFTCAFPPEVVGPKPESPAPPEPPPESEAPDLREERWRAFLALRPTL
RAEVRREARPEVREGQLCLAFFEDKAFHYKASEQKARLLPLAQAHFGVEVVLVIEGE
KKSLSEPPRPAPFPAPFPPEPEVEVEAEAEAEAEAEALRRVRLGGRVLMVWRFP
RTRAEFEEERLQSDSISGTTGI"

```

## ORIGIN

```

1  gtgagcgccc  tctacgcgccc  ottcgcgcccc  ctcacctccc  aggaggtggt  ggggcaggag
61  cactgtgaagg  agccccctcct  caaggcccatc  cgggaggggga  ggctcgccca  ggctacacctc
121  ttctccgggc  ccagggggctt  gggcaagacc  accacggcgg  ggctcctcgc  catggcggtg
181  ggggtgccagg  gggaaagaccc  ccttgccggg  gtctgcccc  actgccaggc  ggtgcagagg
241  ggcgcccacc  cggacgttgt  gtagattgac  gccgcagaca  acaactcctg  ggaggaagtg
301  cgggagctga  gggaaaggat  caacctcgcc  cccctttctg  cccccaggaa  ggtcttcac
361  ctggacgagg  ccacatgctg  ctccaaaagc  gccctaaacg  cctcctcaaa  gacctggag
421  ggcgcccccc  cccacgtcct  ctctgtcttc  gccaccaccg  agcccgagag  gatgcccccc
481  accatctct  cccgcaccca  gcacttcgcg  ttcgcgcggc  tcacggagg  ggagatcgcc
541  tttaagctcc  ggcgcacct  ggaggcgtg  gggcgggagg  cggaggagg  ggccctctc
601  ctctcgcccc  gctcgccgga  cggggccctt  agggacgcgg  aaagcctcct  ggagcgcttc
661  ctctcctg  aaggccccct  cccccgaag  gaggtggagc  gccctcagg  cctcccccc
721  agggaggccc  tggcgagat  cgcgcctcc  ctgcgaggg  ggaanaagcg  gaggccctg
781  ggcctcgccc  ggcgcctcta  cggggaagg  tacgccccc  ggagcctggt  ctggggcctt
841  ttgaggtgt  tccgggaagg  cctctacgcc  gccctggcc  tcgcgggaac  cccctctcc
901  gcccgcccc  aggcctgat  cgcgcctg  accgcctgg  acgagccat  ggagcgctc
961  gccgcgcgt  ccgacgcctt  aagcctggag  gtggccctcc  tggaggcgg  aaggccctg
1021  gccgcgagg  cctgcgccta  gccacgggc  gclcccccc  cagaggtcgg  ccccaagcg
1081  gaaagcccc  cggccccgga  acccccagg  cccgaggagg  ccccccacct  cggggagcgg
1141  tggcgggcct  tctcgaggc  cctcaggccc  accctacggg  cctctgtg  ggagccctg
1201  cgggaggtcc  gggaaaggcca  gctctgctc  gcttccccg  aggcacaagg  ctccacctc
1261  gcgaaggcct  cggaaacagaa  ggcgaggctc  ctcgccctg  cccaggccca  ttctggggtg
1321  gaggaggtcg  ttctgtcct  ggaggggagaa  aaaaaaagc  tgagcccaag  gcccgcccg
1381  gccccaacct  ctgaagcgcc  cgcaccccc  ggccctccc  aggaggaggt  agagcggag
1441  gaagcgggcg  aggagggccc  ggaggaggcc  ttgagcggg  tggtccgct  cctggggggg
1501  cgggtgctct  ggttcggcg  gccaggacc  cgggagcgcc  cggaggagga  accctcgagc
1561  caagacgaga  tagggggtag  tggatataa

```

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Aug 13 2006 12:27:38



```
#####
# Program: needle
# Rundate: Tue Sep 05 13:47:17 2006
# Align_format: srspair
# Report_file: /ebi/extern/old-work/needle-20060905-18471307965295.output
#####
```

```
#####
```

```
#
# Aligned_sequences: 2
# 1: SEQ_3
# 2: Masui
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1594
# Identity: 1581/1594 (99.2%)
# Similarity: 1581/1594 (99.2%)
# Gaps: 8/1594 ( 0.5%)
# Score: 10558.5
#
#
#####
```

SEQ_3	1	gtgagcgccctctaccgcgcttcgcgccctcacctccaggagggtgt	50
Masui	1	gtgagcgccctctaccgcgcttcgcgccctcacctccaggagggtgt	50
SEQ_3	51	ggggcaggagcacgtgaaggagccctctcaaggccatcgggaaggga	100
Masui	51	ggggcaggagcacgtgaaggagccctctcaaggccatcgggaaggga	100
SEQ_3	101	ggctcgccaggcctacctctctcggggccaggggcggtgggcaagacc	150
Masui	101	ggctcgccaggcctacctctctcggggccaggggcggtgggcaagacc	150
SEQ_3	151	accacggcgaggctcctcgccatggcggtgggtgccagggggaagacc	200
Masui	151	accacggcgaggctcctcgccatggcggtgggtgccagggggaagacc	200
SEQ_3	201	cccttgcggggtctgccccactgcacggcggtgcagggggcgccacc	250
Masui	201	cccttgcggggtctgccccactgcacggcggtgcagggggcgccacc	250
SEQ_3	251	cggacgtggtggacattgacgcgcgcagcaacaactccgtggaggacgtg	300
Masui	251	cggacgtggtggagattgacgcgcgcagcaacaactccgtggaggacgtg	300
SEQ_3	301	cggagagctgagggaaaggatccacotcgccccctctctgccccaggaa	350
Masui	301	cggagagctgagggaaaggatccacotcgccccctctctgccccaggaa	350
SEQ_3	351	ggtcttcatcctggacgagggccacatgctctccaaaagcgcttcaacg	400
Masui	351	ggtcttcatcctggacgagggccacatgctctccaaaagcgcttcaacg	400

SEQ_3	401	ccctctctcaagaccctggaggagccccccgccccacgtctcttctgtcttc	450
Masui	401	ccctctctcaagaccctggaggagccccccgccccacgtctcttctgtcttc	450
SEQ_3	451	gccaccacccgagcccgaggagatgccccccaccatctctccccgaccca	500
Masui	451	gccaccacccgagcccgaggagatgccccccaccatctctccccgaccca	500
SEQ_3	501	gcacttccgcttccgcccgtcacggaggaggagatcgccctttaagctcc	550
Masui	501	gcacttccgcttccgcccgtcacggaggaggagatcgccctttaagctcc	550
SEQ_3	551	ggcgcatcctggaggccgtggggcgggaggcgaggaggaggccctctc	600
Masui	551	ggcgcatcctggaggccgtggggcgggaggcgaggaggaggccctctc	600
SEQ_3	601	ctctcgcgccgctggcggaaggcccttagggagcgggaaagcctct	650
Masui	601	ctctcgcgccgctggcggaaggcccttagggagcgggaaagcctct	650
SEQ_3	651	ggagcgcttctctctctggaaggccccctcacccggaaggagtgaggc	700
Masui	651	ggagcgcttctctctctggaaggccccctcacccggaaggagtgaggc	700
SEQ_3	701	gcgccctagg-ctccccccaggga---ccgggtggccgagatcgccgc	746
Masui	701	gcgccctaggcctccccccagggaaggccc---tgcccgagatcgccgc	746
SEQ_3	747	ctccctcgcgagggggaaaaaggcgaggccctgggctcgcccgccgc	796
Masui	747	ctccctcgcgagggggaaaaaggcgaggccctgggctcgcccgccgc	796
SEQ_3	797	tctacggggaagggtacgccccgaggagccttgctcgggccttttgag	846
Masui	797	tctacggggaagggtacgccccgaggagccttgctcgggccttttgag	846
SEQ_3	847	gtgttccgggaaggcctctacgcgcgccttcggcctcgcggaacccccct	896
Masui	847	gtgttccgggaaggcctctacgcgcgccttcggcctcgcggaacccccct	896
SEQ_3	897	tccgcccccgccccaggccctgatcgccgcacatgacccgcctggacgagg	946
Masui	897	tccgcccccgccccaggccctgatcgccgcacatgacccgcctggacgagg	946
SEQ_3	947	ccatggagcgctcgcccgcgctccgacgccttaagcctggagggtggcc	996
Masui	947	ccatggagcgctcgcccgcgctccgacgccttaagcctggagggtggcc	996
SEQ_3	997	ctcctggaggcggggaaggccctggcgccgagccctacccagccac	1046
Masui	997	ctcctggaggcggggaaggccctggcgccgagccctgcccagccac	1046
SEQ_3	1047	ggcgctctctccccagaggtcgcccccaagcgggaaagcccccgacc	1096
Masui	1047	ggcgctctctctccccagaggtcgcccccaagcgggaaagcccccgacc	1096
SEQ_3	1097	cggaacccccaaaggcccgaggagcgcccgacctgcgggagcggtggcg	1146
Masui	1097	cggaacccccaaaggcccgaggagcgcccgacctgcgggagcggtggcg	1146



SEQ_3	1147	gccttctctgaggccctcaggcccccctacgggcttctgtgaggaggg	1196
Masui	1147	gccttctctgaggccctcaggcccccctacgggcttctgtgaggaggg	1196
SEQ_3	1197	ccgccccggaggtccgggaaggccagctctgcctcgtcttccccgaggaca	1246
Masui	1197	ccgccccggaggtccgggaaggccagctctgcctcgtcttccccgaggaca	1246
SEQ_3	1247	aggccttccactaccgcaaggccotcggaacagaaggtgaggtcctcccc	1296
Masui	1247	aggccttccactaccgcaaggccotcggaacagaaggtgaggtcctcccc	1296
SEQ_3	1297	ctggcccaggcccatttcggggtggaggaggtcgtcctcgtcctggaggg	1346
Masui	1297	ctggcccaggcccatttcggggtggaggaggtcgtcctcgtcctggaggg	1346
SEQ_3	1347	agaaaaaaaaagcctgagcccaaggccccgcccgccccacctcctgaag	1396
Masui	1347	agaaaaaaaaagcctgagcccaaggccccgcccgccccacctcctgaag	1396
SEQ_3	1397	cgcccgacccccgggcccctcccgaggaggaggtagaggcggaggaagcg	1446
Masui	1397	cgcccgacccccgggcccctcccgaggaggaggtagaggcggaggaagcg	1446
SEQ_3	1447	gcggaggaggcccccgaggaggccttgaggcgggtggtccgcctcctggg	1496
Masui	1447	gcggaggaggcccccgaggaggccttgaggcgggtggtccgcctcctggg	1496
SEQ_3	1497	ggggcgggtgctctgggtgcggcgccaggacccgggaggcgccggagg	1546
Masui	1497	ggggcgggtgctctgggtgcggcgccaggacccgggaggcgccggagg	1546
SEQ_3	1547	aggaacccctgagccaagacgagatagggggtactcggtatataa	1590
Masui	1547	aggaacccctgagccaagacgagatagggggtactcggtatataa	1590

#####

```
# Program: needle
# Rundate: Tue Sep 05 18:49:39 2006
# Align_format: strapair
# Report_file: /ebi/extern/old-work/needle-20060905-16493852183474.output
#####
```

=====

```
#
# Aligned_sequences: 2
# 1: SEQ_3
# 2: Henne
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1594
# Identity: 1578/1594 (99.0%)
# Similarity: 1578/1594 (99.0%)
# Gaps: 8/1594 ( 0.5%)
# Score: 10532.5
#
#
```

#####

SEQ_3	1	gtgagcgccctctacgcgcgcttcgccccctcaccttcaggaggtgt	50
Henne	1	gtgagcgccctctacgcgcgcttcgccccctcaccttcaggaggtgt	50
SEQ_3	51	ggggcaggagcacgtgaaggagccccctcctcaaggccatccggaggggga	100
Henne	51	ggggcaggagcacgtgaaggagccccctcctcaaggccatccggaggggga	100
SEQ_3	101	ggctcgcccaggcctacctctctcggggccaggggcgtaggcaagacc	150
Henne	101	ggctcgcccaggcctacctctctcggggccaggggcgtaggcaagacc	150
SEQ_3	151	accacggcgaggctcctcgccatggcggtgggtgccagggggaagacc	200
Henne	151	accacggcgaggctcctcgccatggcggtgggtgccagggggaagacc	200
SEQ_3	201	cccttcgggggtctgccccactgccaggcggtgcagaggggcgccacc	250
Henne	201	cccttcgggggtctgccccactgccaggcggtgcagaggggcgccacc	250
SEQ_3	251	cggacgtggtggacattgacgcgcgcagcaacaactcgtggaggacgtg	300
Henne	251	cggacgtggtggagattgacgcgcgcagcaacaactcgtggaggacgtg	300
SEQ_3	301	cgggagctgagggaaaaggatccacctcgccccctctctgccccaggaa	350
Henne	301	cgggagctgagggaaaaggatccacctcgccccctctctgccccaggaa	350
SEQ_3	351	ggtcttcatcctggacgagggcccaatgctctccaaaagcgccctcaacg	400
Henne	351	ggtcttcatcctggacgagggcccaatgctctccaaaagcgccctcaacg	400

SEQ_3	401	ccctcctcaagacccctggaggagcccccccccaagtcctctctgtcttc	450
Henne	401	ccctcctcaagacccctggaggagcccccccccaagtcctctctgtcttc	450
SEQ_3	451	gccaccaccgagcccgagaggatgccccccaccatctctcccgaccca	500
Henne	451	gccaccaccgagcccgagaggatgccccccaccatctctcccgaccca	500
SEQ_3	501	gcacttcgcttccgcgcctcacggaaggaggatcgctttaagctcc	550
Henne	501	gcacttcgcttccgcgcctcacggaaggaggatcgctttaagctcc	550
SEQ_3	551	ggcgcatcctggaggccgtggggcgaggaggaggaggagccctctc	600
Henne	551	ggcgcatcctggaggccgtggggcgaggaggaggaggagccctctc	600
SEQ_3	601	ctctcgcccgctggcggaacggggccttagggagcggaaagcctct	650
Henne	601	ctctcgcccgctggcggaacggggccttagggagcggaaagcctct	650
SEQ_3	651	ggagcgttctctcctcctggaaagccccctcacccggaaggagtgagc	700
Henne	651	ggagcgttctctcctcctggaaagccccctcacccggaaggagtgagc	700
SEQ_3	701	gcgccctagg-ctccccccaggga---ccggggtggcgagatcgccgc	746
Henne	701	gcgccctaggcctccccccagggaagggccc---tgccgagatcgccgc	746
SEQ_3	747	ctcctcgcgagggggaaacggcgaggccctgggctcgccggcgcc	796
Henne	747	ctcctcgcgagggggaaacggcgaggccctgggctcgccggcgcc	796
SEQ_3	797	tctacggggaagggtacgccccgaggagcctggtctcgggccttttgag	846
Henne	797	tctacggggaagggtacgccccgaggagcctggtctcgggccttttgag	846
SEQ_3	847	gtgttccgggaaggcctctacgccgcttcggcctcgcggaacccccct	896
Henne	847	gtgttccgggaaggcctctacgccgcttcggcctcgcggaacccccct	896
SEQ_3	897	tccccccccccccaggccctgatcgccgccatgaccgacctggacgag	946
Henne	897	tccccccccccccaggccctgatcgccgccatgaccgacctggacgag	946
SEQ_3	947	ccatggagcgctcgcccgccgtccgacgcttaagcctggaggtggcc	996
Henne	947	ccatggagcgctcgcccgccgtccgacgcttaagcctggaggtggcc	996
SEQ_3	997	ctcctggaggcggaaggccctggccgcccaggccctacccagccac	1046
Henne	997	ctcctggaggcggaaggccctggccgcccaggccctacccagccac	1046
SEQ_3	1047	ggcgctccttccccagaggtcgggccccaaagccggaagcccccgacc	1096
Henne	1047	ggcgctccccccccagaggtcgggccccaaagccggaagcccccgacc	1096
SEQ_3	1097	cggaaacccccaaaggcccgaggaggcccgacctcgggagcggtggcgg	1146
Henne	1097	cggaaacccccaaaggcccgaggaggcccgacctcgggagcggtggcgg	1146

SEQ_3	1147	gcctctctcgaggccctcaggcccaccctacgggccttcgtgaggaggc	1196
Henne	1147	gcctctctcgaggccctcaggcccaccctacgggccttcgtgaggaggc	1196
SEQ_3	1197	cggcccgagggtcgggaaggccagctctgcctcgtttcccgaggaca	1246
Henne	1197	cggcccgagggtcgggaaggccagctctgcctcgtttcccgaggaca	1246
SEQ_3	1247	aggccttccactacggcaaggcctcggaacagaaggtgaggtcctcccc	1296
Henne	1247	aggccttccactacggcaaggcctcggaacagaaggtgaggtcctcccc	1296
SEQ_3	1297	ctggcccaaggccatttcgggggtgaggagggtcgtcctcgtcctggagg	1346
Henne	1297	ctggcccaaggccatttcgggggtgaggagggtcgtcctcgtcctggagg	1346
SEQ_3	1347	agaaaaaaaaagcctgagcccaaggcccccggcccccacctcctgaag	1396
Henne	1347	agaaaaaaaaagcctgagcccaaggcccccggcccccacctcctgaag	1396
SEQ_3	1397	cgcccgccacccccggggccctcccgaggaggaggtagaggcggaggaaagc	1446
Henne	1397	cgcccgccacccccggggccctcccgaggaggaggtagaggcggaggaaagc	1446
SEQ_3	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
Henne	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
SEQ_3	1497	ggggcgggtgctctgggtgcggcgcccaaggacccgggaggcgccggagg	1546
Henne	1497	ggggcgggtgctctgggtgcggcgcccaaggacccgggaggcgccggagg	1546
SEQ_3	1547	aggaaccctgagccaagacgagataggggtactcggtatataa	1590
Henne	1547	aggaaccctgagccaagacgagataggggtactcggtatataa	1590



# CLUSTAL W (1.83): Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ\_2 529 aa

Sequence 2: SEQ\_4 464 aa

Sequence 3: SEQ\_5 454 aa

Sequence 4: Masui\_tau 529 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 98

Sequences (1:3) Aligned. Score: 99

Sequences (1:4) Aligned. Score: 98

Sequences (2:3) Aligned. Score: 99

Sequences (2:4) Aligned. Score: 96

Sequences (3:4) Aligned. Score: 97

Guide tree file created: [/ebi/extserv/clustalw-

work/interactive/clustalw-20060905-18544611.dnd]

Start of Multiple Alignment

There are 3 groups

Aligning...

Group 1: Sequences: 2 Score:11312

Group 2: Sequences: 2 Score:9756

Group 3: Sequences: 4 Score:9812

Alignment Score 17218

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060905-18544611.aln]

## CLUSTAL W (1.83) multiple sequence alignment

```

SEQ_2      MSALYRRFRRLTFTQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
Masui_tau  MSALYRRFRRLTFTQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_4      MSALYRRFRRLTFTQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_5      MSALYRRFRRLTFTQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
*****

SEQ_2      GCQGEDPPCGVCPHCQAVQRGAFHDVVDDAASNNNSVEDVRELRLRIHLAPLSAPRKVFI 120
Masui_tau  GCQGEDPPCGVCPHCQAVQRGAFHDVVDDAASNNNSVEDVRELRLRIHLAPLSAPRKVFI 120
SEQ_4      GCQGEDPPCGVCPHCQAVQRGAFHDVVDDAASNNNSVEDVRELRLRIHLAPLSAPRKVFI 120
SEQ_5      GCQGEDPPCGVCPHCQAVQRGAFHDVVDDAASNNNSVEDVRELRLRIHLAPLSAPRKVFI 120
*****

SEQ_2      LDEAHMLSKSAFNALLKLTLEPPPHVLFVFATTEPERMPFTILSRQHFRRRLTEEFIA 180
Masui_tau  LDEAHMLSKSAFNALLKLTLEPPPHVLFVFATTEPERMPFTILSRQHFRRRLTEEFIA 180
SEQ_4      LDEAHMLSKSAFNALLKLTLEPPPHVLFVFATTEPERMPFTILSRQHFRRRLTEEFIA 180
SEQ_5      LDEAHMLSKSAFNALLKLTLEPPPHVLFVFATTEPERMPFTILSRQHFRRRLTEEFIA 180
*****

SEQ_2      FKLRRLLEAVGREAEAEALLLLARLADGALRDAESLLERFLLEGLPLTRKEVERALGSPF 240
Masui_tau  FKLRRLLEAVGREAEAEALLLLARLADGALRDAESLLERFLLEGLPLTRKEVERALGSPF 240
SEQ_4      FKLRRLLEAVGREAEAEALLLLARLADGALRDAESLLERFLLEGLPLTRKEVERALGSPF 240
SEQ_5      FKLRRLLEAVGREAEAEALLLLARLADGALRDAESLLERFLLEGLPLTRKEVERALGSPF 240
*****

SEQ_2      GTGVAEIAASLARGKTAELGLARALYEGGYAPRSLSVGLLEVFREGLYAFLAGTFLP 300

```

```

Masui_tau      REALAEIRASLARGKTAELGLARRLYGEGYAPRSLVSGLELVFREGLYAAFGLAGTPLP 300
SEQ_4          GTGVANIAASLARGKTAELGLAPRLYGEGYAPRSLVSGLELVFREGLYAAFGLAGTPLP 300
SEQ_5          GTGVAFYIASLARGKTAELGLARRLYGEGYAPRSLVSGLELVFREGLYAAFGLAGTPLP 300
                :*****

```

```

SEQ_2          APPQALIAAMTALDEAMERLARSSDALSLEVALLEAGRALAAEALPQPTGAPSFVEVGPKP 360
Masui_tau      APPQALIAAMTALDEAMERLARSSDALSLEVALLEAGRALAAEALPQPTGAPSFVEVGPKP 360
SEQ_4          APPQALIAAMTALDEAMERLARSSDALSLEVALLEAGRALAAEALPQPTGAPSFVEVGPKP 360
SEQ_5          APPQALIAAMTALDEAMERLARSSDALSLEVALLEAGRALAAEALPQPTGAPSFVEVGPKP 360
                *****

```

```

SEQ_2          ESPPTPEPPRFEELPDLRERWRFLALRPTLRAFVREARPEVREGQLCLAFFEDKAFHY 420
Masui_tau      ESPPAPEPPRFEELPDLRERWRFLALRPTLRAFVREARPEVREGQLCLAFFEDKAFHY 420
SEQ_4          ESPPTPEPPRFEELPDLRERWRFLALRPTLRAFVREARPEVREGQLCLAFFEDKAFHY 420
SEQ_5          ESPPTPEPPRFEELPDLRERWRFLALRPTLRAFVREARPEVREGQLCLAFFEDKAFHY 420
                *****

```

```

SEQ_2          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKKSLSPPRPAPPPPEAPAPPGPPEEEVEAE 480
Masui_tau      RKASEQKARLLPLAQAHFGVEEVVLVLEGEKKSLSPPRPAPPPPEAPAPPGPPEEEVEAE 480
SEQ_4          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KPEPKAPPGPTS----- 464
SEQ_5          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KA----- 454
                *****

```

```

SEQ_2          EAAEEAPEEALRRVVLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
Masui_tau      EAAEEAPEEALRRVVLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
SEQ_4          -----
SEQ_5          -----

```

# CLUSTAL W (1.82) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ\_2 529 aa

Sequence 2: SEQ\_4 464 aa

Sequence 3: SEQ\_5 454 aa

Sequence 4: Henne\_tau 529 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 98

Sequences (1:3) Aligned. Score: 99

Sequences (1:4) Aligned. Score: 98

Sequences (2:3) Aligned. Score: 99

Sequences (2:4) Aligned. Score: 96

Sequences (3:4) Aligned. Score: 97

Guide tree file created: [/ebi/extserv/clusterw-

work/interactive/clusterw-20060905-18580167.dnd]

Start of Multiple Alignment

There are 3 groups

Aligning...

Group 1: Sequences: 2 Score:11303

Group 2: Sequences: 2 Score:9756

Group 3: Sequences: 4 Score:9802

Alignment Score 17203

CLUSTAL-Alignment file created [/ebi/extserv/clusterw-work/interactive/clusterw-20060905-18580167.aaln]

## CLUSTAL W (1.83) multiple sequence alignment

```

SEQ_2      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFGSPRGVGKTTTARLLAMAV 60
Henne_tau MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFGSPRGVGKTTTARLLAMAV 60
SEQ_4      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFGSPRGVGKTTTARLLAMAV 60
SEQ_5      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFGSPRGVGKTTTARLLAMAV 60
*****

```

```

SEQ_2      GCQGEDPFCGVCPHCQAVQRGAPDVVDIDAASNNNSVEDVRELREIRHLAPLSAPFRKVF 120
Henne_tau GCQGEDPFCGVCPHCQAVQRGAPDVVDIDAASNNNSVEDVRELREIRHLAPLSAPFRKVF 120
SEQ_4      GCQGEDPFCGVCPHCQAVQRGAPDVVDIDAASNNNSVEDVRELREIRHLAPLSAPFRKVF 120
SEQ_5      GCQGEDPFCGVCPHCQAVQRGAPDVVDIDAASNNNSVEDVRELREIRHLAPLSAPFRKVF 120
*****

```

```

SEQ_2      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTOHFRFRRLTEEEIA 180
Henne_tau LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTOHFRFRRLTEEEIA 180
SEQ_4      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTOHFRFRRLTEEEIA 180
SEQ_5      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTOHFRFRRLTEEEIA 180
*****

```

```

SEQ_2      FKLRLILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSP 240
Henne_tau FKLRLILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSP 240
SEQ_4      FKLRLILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSP 240
SEQ_5      FKLRLILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSP 240
*****

```

```

SEQ_2      GTGVAEIAASLARGKTAEALGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTFLP 300
Henne_tau REALAEIAASLARGKTAEALGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTFLP 300

```



```

SEQ_4      GTGVAETIAASLARGKTAEALGLARRLYGEGYAPRSLVSSLLLEVFRGLVAAFGLAGTFLP 300
SEQ_5      GTGVAETIAASLARGKTAEALGLARRLYGEGYAPRSLVSSLLLEVFRGLVAAFGLAGTFLP 300
           .:*****

SEQ_2      APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPFVGPKP 360
Henne_tau APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPFVGPKP 360
SEQ_4      APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPFVGPKP 360
SEQ_5      APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPFVGPKP 360
           *****

SEQ_2      ESPPTFEPFPPEEAPDLRERWRAPFLEALRPTLRAPVREARPEVREGQLCLAFPEDKAFHY 420
Henne_tau ESPPTFEPFPPEEAPDLRERWRAPFLEALRPTLRAPVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_4      ESPPTFEPFPPEEAPDLRERWRAPFLEALRPTLRAPVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_5      ESPPTFEPFPPEEAPDLRERWRAPFLEALRPTLRAPVREARPEVREGQLCLAFPEDKAFHY 420
           *****

SEQ_2      RKASEQKVRLPLAQAHFGVEEVVLVLEGEKKSLSPRFRFAPFPPEAPAPPGFPPEEEVEAE 480
Henne_tau RKASEQKVRLPLAQAHFGVEEVVLVLEGEKKSLSPRFRFAPFPPEAPAPPGFPPEEEVEAE 480
SEQ_4      RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KPEPKAPPGPTS----- 464
SEQ_5      RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KA----- 454
           *****

SEQ_2      EAAEEAFEEALRRVVRLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
Henne_tau EAAEEAFEEALRRVVRLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
SEQ_4      -----
SEQ_5      -----

```